Administrative Boundaries: Automating the Data Processing Cycle for a Critical National Dataset

Mr Michael Dixon
Products Business Manager
PSMA Australia

Mr Tom Spencer
Technical Consultant
1Spatial

PSMA Australia Limited

PSMA Australia Limited is an unlisted public company, established under Australia’s Corporations Act, wholly owned by the State, Territory and Australian Governments. The organisation provides an important bridge between data custodians and solution providers.
Reference Datasets

Transport & Topography
Over 2,000,000 kms of named road centrelines in a structured hierarchy maintained quarterly. National rail network including tram lines. Airports and Landing Grounds. National Drainage network consisting of Major and Minor water layers and polygon water bodies.

CadLite
Over 10.8 million polygons representing the registered land parcels in Australia updated quarterly with incremental updates available. Every parcel contains the legal parcel identifier that acts as a key to access richly attributed jurisdictional Digital Cadastral Data Bases (DCDB). Also contains links to key administrative data layers including Local Government Area and Locality. A property version of the dataset is also available.

G-NAF
G-NAF contains over 12 million physical addresses and approximately 2.5 million aliases updated quarterly. Data is sourced from AEC, Australia Post and Government Mapping Agencies and Land Registries. Every address contains a Geocode (Latitude & Longitude) and metadata to assist in decision making. Sophisticated data modelling to enhance application accuracy.

Reference Datasets

Points of Interest
Over 180,000 Points of Interest including:
• Police Stations
• Hospitals
• Post Offices
• Museums
• Churches
• Airports
• Banks
• Swimming pools
• Libraries
• Theatres
• Shopping Centres

Postal Boundaries
This definitive dataset has been developed by Australia Post and PSMA Australia and is updated quarterly. Includes two layers:
• Boundaries – polygon data
• Centroid – point data

Administrative Boundaries
This dataset contains all of Australia’s major administrative boundaries including:
• Key ABS Statistical Geography
  • Mesh Blocks
  • Collector districts
  • Statistical local areas
• Urban centres/localities
• State Boundaries
• Electoral Boundaries
  • Commonwealth
  • State and Territory
• Local Government Areas
• Suburbs/localities
• Town points
The Case for Change

- Time to Market: Current PSMA Australia release cycle is quarterly.
  - Hierarchical processing required, dependency between datasets
  - Downstream (VAR) processes must be added – further delay
  - Processes are not homogeneous: Multiple Data Managers, Work Practices, Standards
  - No direct involvement in processing: Over time PSMA Australia was moving further away from understanding the specific processes involved.
  - Lack of transparency: Difficult to have confidence in repeatability; particularly in atypical situations.
  - Innovation: No incentive to innovate beyond your sphere of influence.
  - Metadata: Only at a dataset level and not dynamic.

The Data Management Cycle

The PSMA Australia data management cycle can be represented as a series of processing steps completed by a number of different actors. Including PSMA Australia, a Data Manager and the LYNX Manager.
Why Radius Studio?

1Spatial’s Radius Studio was selected to provide the “data management” components of the supply chain, specifically:
- Conformance to Rules
- Standardisation

Key factors behind the decision:
- Built on SOA
- Rich range of functions including:
  - Spatial Rules Engine
  - Metadata Recording
  - Strong Reporting Capability

Business Rules behind the data maintenance cycle needed to be clearly understood. Major benefit for PSMA Australia.

Administrative Boundaries: Business Rules

These are the fundamental rules that the data in each dataset must conform to. Typically they are quite logical and should be relatively easy to understand. e.g. adjoining suburbs should not overlap.

Each rule is created as a new service inside Radius Studio which can be reused across different contributor supplies, themes etc.
The Challenges

- Automation of Administrative Boundaries Updates
  - Consistent inputs
  - What happens if...
  - Fit within larger data management process

- Minimise duplication of functionality between Contributors and also between Datasets

- Re-usable data management services

- Internal knowledge of datasets & software
### Technical Implementation

<table>
<thead>
<tr>
<th>Mapping</th>
<th>Cleansing</th>
<th>Updating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Facilitate reuse of functionality</td>
<td>Express product specification</td>
<td>Manage feature lifecycle</td>
</tr>
<tr>
<td>Generic Schema</td>
<td>Measurable data quality</td>
<td>Recording changes</td>
</tr>
<tr>
<td>Consistent inputs</td>
<td>Enforce product specification</td>
<td></td>
</tr>
</tbody>
</table>

### Administrative Boundaries Data Mapping

<table>
<thead>
<tr>
<th>Mapping</th>
<th>Cleansing</th>
<th>Updating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Facilitate reuse of functionality</td>
<td>Express product specification</td>
<td>Manage feature lifecycle</td>
</tr>
<tr>
<td>Generic Schema</td>
<td>Measurable data quality</td>
<td>Recording changes</td>
</tr>
<tr>
<td>Consistent inputs</td>
<td>Enforce product specification</td>
<td></td>
</tr>
</tbody>
</table>
Technical Implementation

Data Mapping

Mapping contributor data onto a generic processing schema facilitates functionality reuse across contributors.

Mapping PSMA data onto a generic processing schema facilitates functionality reuse across datasets.

Technical Implementation

Data Cleansing

Mapping
- Facilitate reuse of functionality
- Generic Schema
- Consistent inputs

Cleansing
- Express product specification
- Measurable data quality
- Enforce product specification

Updating
- Manage feature lifecycle
- Recording changes
Technical Implementation
Data Cleansing

BR11 – Polygons will not overlap

"The polygon with the largest area will have the area of overlap removed"
**Administrative Boundaries Data Update**

**Mapping**
- Facilitate reuse of functionality
- Generic Schema
- Consistent inputs

**Cleansing**
- Express product Specification
- Measurable data quality
- Enforce product specification

**Updating**
- Manage feature lifecycle
- Record changes

---

**Technical Implementation Data Update**

[Diagram showing technical implementation process]
The Benefits for PSMA

- Increased understanding of its core business.
- Continued provision of high quality datasets at an increased frequency.
- Increased flexibility to adapt to changing market conditions.
- Increased data quality.
- PSMA Australia to have greater control of the processes behind the data creation.
- Standardised data maintenance practices.
- Capability to accept data from non traditional sources.
- Closer relationship with Jurisdictions and other data providers.
- Less labour intensive data maintenance processes.
- International relationships through collaboration on solving continental scale data management challenges.

Questions

Michael Dixon
Products Business Manager

Tom Spencer
Technical Consultant

Answers

Contact PSMA Australia
+61 2 6260 9000
enquiries@psma.com.au
www.pisma.com.au

Contact 1Spatial Australia
+61 2 6260 9081
info@1spatial.com
www.1spatial.com